DOCUMENT RESUME

ED 414 327 TM 027 853

AUTHOR Foote, Chandra J.
TITLE Item-Specific Gender Differences in Confidence.
PUB DATE 1997-00-00
NOTE 19p.
PUB TYPE Reports - Research (143)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Educational Research; *Graduate Students; Higher Education; Problem Solving; *Sex Differences; *Student Attitudes; *Test Items; *Undergraduate Students
IDENTIFIERS Accuracy; Calibration; *Confidence

ABSTRACT

Very little research has been performed which examines gender differences in confidence in highly specified situations. More generalized studies consistently suggest that women are less confident than men (i.e. Sadker and Sadker, 1994). The few studies of gender differences in item-specific conditions indicate that men tend to be more confident in their responses to more ambiguous, problem-solving questions than women, and women are at least as confident in answering less ambiguous, factual-level questions (i.e. Lundeberg et al, 1994). This research also indicates that undergraduate men are not as well-calibrated to assess accuracy as women and more experienced men. The present study examines item-specific confidence using students enrolled in two sections of an educational psychology course. Twenty-six students in the in-service section have their bachelors degree in education and are returning for their masters degree for permanent certification. Thirty-three students in the pre-service section have their bachelors degree in another field and are returning for their masters degree in education. The results reveal no significant gender or experience level differences in accuracy, confidence, or calibrations. Implications and future avenues of research are discussed. (Contains 4 tables and 14 references.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 414 327

Item-Specific Gender Differences in Confidence

Chandra J. Foote, Niagara University

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL
HAS BEEN GRANTED BY
Chandra Foote

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☑ This document has been reproduced as
received from the person or organization
originating it.
☐ Minor changes have been made to
improve reproduction quality.
• Points of view or opinions stated in this
document do not necessarily represent
official OERI position or policy.

Running Head: Gender Differences in Confidence

TMO27853

BEST COPY AVAILABLE

Abstract

Very little research has been performed which examines gender differences in confidence in highly specified situations. More generalized studies consistently suggest that women are less confident than men (i.e. Sadker & Sadker, 1994). The few studies of gender differences in item-specific confidence indicate that men tend to be more confident in their responses to more ambiguous, problem-solving questions than women, and women are at least as confident in answering less ambiguous, factual level questions (i.e. Lundeberg et al. 1994). This research also indicates that undergraduate men are not as well-calibrated to assess accuracy as women and more experienced men. The present study examines item-specific confidence using students enrolled in two sections of an educational psychology course. The results reveal no significant gender or experience level differences in accuracy, confidence, or calibrations. Implications and future avenues of research are discussed.

Item-Specific Gender Differences in Confidence

Confidence in personal ability has often been suggested as a factor that motivates men to persist in professional fields such as science and engineering, while less confident women avoid these fields (i.e. Dix, 1987). With little exception, studies of gender and motivation have reported that females from third grade through adulthood hold lower expectations for success and are less likely to attribute success to ability than males (Licht, Stader & Swenson, 1989; Sadker & Sadker, 1994; and Schunk & Lilly, 1984). Clance and O'Toole (1988) studied successful professional women and found that they tended to underestimate their own abilities while overestimating those of others. This became known as the "Imposter Phenomenon" because women lacked confidence in their personal ability and felt that their successful positions were a misrepresentation.

Most studies of gender differences in confidence take a generalized approach asking subjects to take an attitude scale, predict grades, or assess their ability to pass a test (Campbell & Hackett, 1986, Hyde, Fennema, Ryan, Frost & Hopp, 1990). Very few studies have taken an item-specific approach to assessing confidence. Lenney (1977) reported that females are more likely to negatively assess their ability when a task is ambiguous or unfamiliar, perhaps due to a socialization process in which girls are taught to hold a lower tolerance for ambiguity. The studies which assess confidence on a general level, such as prediction of grade or ability to pass a test, are potentially more

ambiguous than studies which make use of item-specific techniques. If an individual is assessing their ability to perform a very specific task such as answering individual test items, then the ambiguous nature of the task is reduced.

Item-specific confidence is analogous to numerous constructs in the meta-cognitive domain of cognitive psychology. Students who rate their confidence in correctly answering a certain question are in essence testing their comprehension monitoring and feeling-of-knowing (i.e. Epstein, Glenberg & Bradley, 1984). It is assumed that more successful individuals would make use of better meta-cognitive strategies by assessing their confidence as higher when they are correct than when they are incorrect.

Two of the rare published studies of item-specific confidence have resulted in similar conclusions. Jones and Jones (1989) asked high school students to assess their ability to answer four questions: two mathematics questions and two science questions. They found that females were less confident on the science questions and math questions that were unfamiliar problem-solving type questions. Yet they were more confident on familiar computational-type math questions. This finding is consistent with Lenney's (1977) review which suggested that women are confident in familiar, unambiguous situations.

Lundeberg, Fox and Puncochar (1994) improved upon the Jones and Jones (1989) study by increasing the number of items assessed by their subjects. They also tested students enrolled in three college level courses on items that were part of the course

work, rather than in an experimental context. Lastly, they had students estimate their confidence after they had actually answered the items. Glenberg and Epstein (1987) demonstrated that subjects were better at monitoring their comprehension abilities after solving a problem rather than before.

The Lundeberg et al. (1994) study provided evidence that males in laboratory methods courses were more confident in their answers when correct and when incorrect than females. There were, however, no significant gender differences in confidence on questions items in a memory course. The test items in a typical laboratory course are likely to be more ambiguous and unfamiliar than items in a memory course. Therefore, the results of this study also support the findings of Lenney (1977) regarding women's confidence on ambiguous items. The Lundeberg et al. (1994) study also revealed that undergraduate men in the memory course were over-confident in their accuracy predictions on incorrect answers. Women and graduate men appeared to be better calibrated at rating their confidence than undergraduate men.

In sum, the findings on item-specific confidence indicate that men tend to be more confident in their responses to more ambiguous, problem-solving questions than women, and women are at least as confident in answering less ambiguous, factual level questions. The research also indicates that undergraduate men are not as well-calibrated to assess accuracy as women and more experienced men.

The present research examines gender differences in confidence on exam items in an education course. The following questions will guide the study: (1) are women more

confident than men on unambiguous exam items in an education course, (2) are students with more teaching experience, more confident than students with less teaching experience, and (3) are females and more experienced graduate students better calibrated to assess accuracy.

Confidence was assessed for students enrolled in two sections of a human learning course taught by the same instructor, using the same text, lesson plans, and method of presentation. It was expected that women would be as confident as men at answering the questions in this course since the material is less ambiguous than many math and science, problem-solving courses, and it parallels much of the material typically taught in a memory course such as that of Lundeberg et. al. (1994) where no gender differences in confidence were found.

Students in the two sections of human learning are in different teacher preparation programs. Students in both sections have received their bachelors degree. However, students in one section have their degree in education and have had extensive field experience, including student teaching and for some, years of classroom teaching. Students in the other section have their bachelors degree in a general liberal arts area and are returning for their masters degree in education. Students in both sections are of similar age and life experience, but the students with a bachelors degree in education have more teaching experience. It was thought that these students would be less likely to consider test questions as ambiguous since they have encountered similar information in their undergraduate course work and professional practice. It was expected that the

students with more teaching experience would be more confident of the information on the test.

Lastly, it was expected that females at both levels of teacher preparation and males with more field experience would be better calibrated to assess confidence than pre-service (less experienced) men. That is, they would be more confident in correct answers than incorrect answers. This result was expected based on the findings of Lundeberg et al. (1994) where undergraduate men were less well-calibrated than women and graduate men.

Predictions:

1. Women were expected to be as confident as men, since the course was a offered to students in a female dominated field, education, and the test items include straight forward, unambiguous questions.

2. More experienced students were expected to be more confident than less experienced students, since the in-service students were more familiar with the information.

3. Women, and male graduate students with more in-service experience were expected to be better calibrated to assess accuracy.

Method

Subjects

Students in two sections of an educational psychology course provided the item-specific confidence ratings. The 26 students in the in-service section have their bachelors degree in education and are returning for their masters degree for permanent certification to teach in New York State. In this section their were 7 men and 19 women. The 33 students in the pre-service section have their bachelors degree in another field and are returning for their masters degree in education. In this section there were 15 men and 18 women. The average age of students in both sections is about 28 years.

Procedure

Students in both sections of the course took a 50-item, multiple-choice test as part of the term requirement. Students were asked to indicate their confidence that an answer was correct, after answering each item. All students were told that their confidence judgments would have no bearing on their course grade, and that they should be as accurate in their judgments as possible. They were offered extra course credit for providing their confidence judgments. All students elected to participate and indicated their confidence on a 5-point Likert scale, with 1=pure guess/totally uncertain, 2=mostly uncertain, 3=mixed feelings of confidence and uncertainty, 4=somewhat confident, 5= totally confident that the answer is correct.

Results

Accuracy

Table 1 presents the mean percent accuracy scores of men and women in each section and for the two sections combined. There were no gender differences in mean accuracy: $t_{(31)}=1.32$, $p=.196$, for the pre-service section; $t_{(24)}=1.59$, $p=.125$, for the in-service section; and $t_{(57)}=.038$ and $p=.97$, for the two sections combined. There were also no significant differences in accuracy between the different sections: $t_{(20)}=2.03$, $p=.055$, for men; $t_{(35)}=.82$, $p=.419$, for women; and $t_{(57)}=.487$, $p=.628$ for men and women combined within each section.

TABLE 1

Mean Percent Accuracy of Women and Men

Sample	n	M	SD
Pre-service			
Men	15	81.1	8
Women	18	76.7	10.6
In-service			
Men	7	71.4	14.4
Women	19	79.5	10.3
Combined			
Men	22	78.0	11.1
Women	37	78.1	10.4

Similarities and Differences in Confidence

Table 2 presents the mean confidence of women and men when they answer correctly and when they answer incorrectly in each section and as combined over both sections. The mean confidence ratings for both sections and the combined mean

confidence ratings resulted in women rating themselves slightly more confident than men. However, tests failed to yield any significant differences in confidence between men and women in either section or when combined (all p's>.10). There were also no significant differences in confidence when correct or incorrect between sections for men, women, or men and women combined within each section (all p's >.40).

TABLE 2

Mean Confidence of Men and Women

		Confidence Correct		Confidence Incorrect	
class	n	M	SD	M	SD
Pre-service					
Men	15	3.80	.54	2.89	.68
Women	18	4.06	.56	2.99	.79
In-Service					
Men	7	3.88	.41	2.86	.36
Women	19	4.00	.51	3.16	.66
Combined					
Men	22	3.82	.50	2.88	.59
Women	37	4.03	.53	3.08	.72

Calibration

A general measure of confidence calibration would be indicated when confidence in correct answers is higher than confidence in incorrect answers. Men and women in each class and overall showed significantly higher confidence ratings when correct than when incorrect: $t_{(14)}=5.67$, $p<.0001$ for pre-service men; $t_{(17)}=6.28$, $p<.0001$ for pre-service women; $t_{(6)}=8.54$, $p<.0001$ for in-service men; $t_{(18)}=7.58$, $p<.0001$ for in-service

women; $t_{(21)}=8.25$, $p<.00001$ for men overall; and $t_{(36)}=9.44$, $p<.00001$ for women overall.

A more specific test of calibration might make use of traditional rank correlation methods. These methods however, would not be suitable for the present data since the variables of confidence and accuracy are tied to individual cases. This type of correlation would eliminate tied rankings. For this reason calibration was examined more in-depth using the Confidence Accuracy Quotient or CAQ (Shaughnessy, 1979). The CAQ method deletes only those subjects who provide all correct or all incorrect responses. There were no cases in which this occurred, therefore the following calibration analysis includes all of the data.

The CAQ index uses the following formula to standardize confidence ratings based on accuracy:

$$CAQ = \frac{\text{M confidence correct-M confidence incorrect}}{\text{SD overall confidence}}$$

Positive CAQ scores indicate better calibration since the individual has higher confidence when correct than when incorrect. Lower or negative CAQ scores indicate poor calibration since the individual's confidence is too high when incorrect.

TABLE 3
Mean Confidence Accuracy Quotient (CAQ) Scores

Course	n	M	SD
Pre-service			
Men	15	.75	.41
Women	18	1.02	.63
Overall	26	.89	.55
In-service			
Men	7	.89	.24
Women	19	.78	.39
Overall	26	.81	.36
Combined			
Men	22	.79	.37
Women	37	.90	.53
Overall	59	.86	.47

Table 3 presents the mean CAQ scores for men and women within each section, mean overall scores within each section, and mean scores for the sections combined. Gender differences in CAQ scores were assessed between men and women within each section, and in the combined scores of the two section. No significant differences were found (p's all > .16). CAQ score differences between classes, for men, women, and men and women combined within sections, were also assessed to determine if the more experienced in-service subjects were better calibrated in their confidence. Again no significant difference was found (p's>.40).

Domain Specific Calibration

It has been suggested that aggregating confidence accuracy scores across an entire test might conceal gender differences that are apparent when tests are broken down into smaller content domains (Lundeberg, Fox, and Puncochar, 1994). A factor analysis was

performed on the 50 multiple choice items on the test to determine if specific content areas or question types (i.e. factual, analytical, evaluation) were evident. This analysis yielded only one factor. Therefore the exam could not be broken into component parts and gender and experience level differences on subsets of items could not be assessed.

Confidence and Accuracy Comparison

Overall confidence scores were derived by taking the mean confidence scores as assessed on the Likert scale and dividing the mean confidence score of each subject by 5, the highest possible confidence rating. As Table 4 illustrates both men and women at the pre-service and in-service levels are under-confident in the accuracy of their answers, with confidence estimates ranging from 4 to 14 points lower than the actual percent correct.

TABLE 4

Mean Accuracy and Confidence

sample	n	accuracy		overall confidence	
		M	SD	M	SD
Pre-service					
Men	15	81.1	8.0	66.8	10.7
Women	18	76.7	10.6	70.5	11.7
In-Service					
Men	7	71.4	14.4	67.3	7.0
Women	19	79.5	10.3	71.6	10.7

Discussion

It was expected that women would be as confident as men when making item-specific judgments. Lenney (1977) indicated that women assess their abilities as lower when a task is ambiguous, and are more confident in well defined situations. The mean confidence results presented in Table 2 demonstrate that women consistently rated their confidence higher than men. The analysis of gender differences in confidence, however, yielded no significant differences. In this course women were as confident as men in their confindence in accuracy judgments.

Students with more teaching experience were expected to be more confident than less experienced students based on their background in education. This expectation was not supported. Students in both sections of the course were equally as confident in their answers.

Women and men with more in-service experience were expected to be better calibrated to assess accuracy than pre-service males. The analysis revealed that all of the students were calibrated to assess their confidence when correct, significantly higher than their confidence when incorrect. The analysis on confidence accuracy quotients failed to reveal any significant differences between men and women or between individuals with differing levels of experience. Specifically, the pre-service men were not over confident in their incorrect answers as had been expected. In fact, in this study all of the students were under confident in their perceptions of accuracy, which directly contradicts past

findings in which students were found to consistently assess their confidence higher than their actual accuracy level (i.e. Lundeberg et al. 1994).

The current study is limited by the type of information tested within the course. This information tended to be very factual and unambiguous. The type of the information tested in this course may also explain why no differences were found between accuracy or confidence estimates of more and less experienced students. Teachers who have been working in the field may be more adept at and confident in ambiguous problem solving situations pertaining to student learning. The factual nature of the current exam items may not have uncovered any differences in accuracy or confidence because they were so clear cut and could be learned from the course instruction alone without experience in field application.

This study provides more evidence that women are equally as confident as men when item-specific judgments are made on unambiguous questions. This study does not however, examine confidence in ambiguous situations. Women appear to be comfortable as educators and tend to enter the field in greater numbers than men. If these women also possess less tolerance for ambiguity, as Lenney (1977) suggests, this may partially explain why society is displeased with the lack of problem-solving ability displayed by the current generation of students. Teachers may avoid modeling problem-solving strategies to their students because they are themselves uncomfortable. Thus, perpetuating students' feelings of confidence and comfort in unambiguous situations and

intolerance for ambiguity. Future research should seek to determine if women are confident on ambiguous questions in female dominated domains like education.

The current study also suggests that less experienced men may not be over-confident when incorrect in all situations. Confidence calibration may be adjusted according to context. In the field of education, where females tend to be socially accepted and confident, less experienced men lacked the over-confidence evident in earlier studies using similar material (i.e. human learning and memory, Lundeberg et al., 1994). An unfortunate contradictory explanation for this finding would be that males with less tolerence for ambiguity are self-selecting teaching as a career. Further research should also be performed to assess the over-confidence issue.

References

Campbell, N.D. & Hackett, G. (1986). The effects of mathematics task performance on math self-efficacy and task interest. Journal of Vocational Behavior, 28, 149-162.

Clance, P. & O'Toole, M. (1988). The impostor phenomenon: An internal barrier to empowerment and achievement. Women and Therapy, 6, 51-64.

Dix , L.S. (Ed.). (1987) Women: Their underrepresentation and career differentials in science and engineering. Proceedings of a workshop at the National Academy of Sciences. Washington, DC: National Academy Press.

Epstein, W., Glenberg, A. & Bradley, M. (1984). Coactivation and comprehension: Contributions of text variables to the illusion of knowing. Memory & Cognition, 12, 355-360

Glenberg, A.M. & Epstein, W. (1987). Inexpert calibration of comprehension. Memory & Cognition, 15, 84-93.

Hyde, J.S., Fennema, E., Ryan, M., Frost, L.A. & Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect: A meta-analysis. Psychology of Women Quarterly, 14, 299-324.

Jones, L.P. & Jones, L.G. (1989). Context, confidence and the able girl. Educational Research, 31, 189-194.

Kramer, L. (1991). The social construction of ability perceptions: An ethnographic study of gifted adolescent girls. Journal of Early Adolescence, 11, 340-362.

Lenney, E. (1977). Women's self-confidence in achievement settings. Psychological Bulletin, 84, 1-13.

Licht, B., Stader, S. & Swenson, C. (1989). Children's achievement-related beliefs: Effects of academic area, sex, and achievement level. Journal of Educational Research, 82, 253-260.

Lundeberg, M.A., Fox, P.W. & Puncochar, J. (1994). Highly confident but wrong: Gender differences and similarities in confidence judgments. Journal of Educational Psychology, 86, 114-121.

Sadker, M. & Sadker, D. (1994). Failing at Fairness: How Schools Cheat Girls. New York: Touchstone.

Schunk, D. & Lilly, M. (1984). Sex differences in self-efficacy and attributions: Influence of performance feedback. Journal of Early Adolescence, 4, 203-213.

Shaughnessy, J. (1979). Confidence-judgement accuracy as a predictor of test performance. Journal of Research in Personality, 13, 505-514.